

11) Publication number:

0 079 739 A2

(12)

EUROPEAN PATENT APPLICATION

- (1) Application number: 82305926.6
- 22 Date of filing: 08.11.82

(i) Int. Cl.³: C 12 N 15/00 C 12 N 1/00, C 12 P 21/02 C 07 H 21/04, C 07 C 103/52 //C12R1/19, C12R1/865

- 30 Priority: 12.11.81 US 320632
- Date of publication of application: 25,05,83 Bulletin 83/21
- Designated Contracting States:
 BE CH DE FR GB IT LI NL SE

- (1) Applicant: THE UPJOHN COMPANY 301 Henrietta Street Kalamazoo, Michigan 49001(US)
- 72 Inventor: Dugaiczyk, Achilles c/o The Upjohn Company 301 Henrietta Street Kalamazoo Michigan 49001(US)
- (24) Representative: Perry, Robert Edward et al, GILL JENNINGS & EVERY 53-64 Chancery Lana London WC2A 1HN(GB)

(54) Albumin-based nucleotides, their replication and use, and plasmids for use therein.

(E) The DNA sequence coding for human serum albumin has been isolated and inserted as two fragments into two novel plasmids which can be replicated in E. colf. These novel fragments can be gioned to provide a unitary DNA sequence which then can be cloned into a suitable host, eq., E. colf. for the expression of human serum albumin (which is used extensively in medical practice in treating shock conditions).

GJE 70/2056/02

ALBUMIN-BASED NUCLEOTIDES, THEIR REPLICATION AND USE, AND PLASMIDS FOR USE THEREIN

This invention relates to nucleotides related to human serum albumin (HSA), their replication and use, and plasmids (and host substances) for use therein.

The gene for serum albumin is regulated in 5 development. On the other hand, serum albumin is synthesised in mammals by the adult liver, and its plateau in adulthood. The embryonic liver and yolk sac, on the other hand, produce predominantly \(\alpha \)-fetoprotein, but the synthesis decreases drastically after birth. Recently, 10 Law et al determined the complete sequence of mouse \(\alpha \)-fetoprotein mRNA, Nature 291 (1981) 201-205. The structure revealed extensive homology to mammalian serum albumin, indicating that the two proteins are encoded in the same gene family. Similar conclusions have been 15 reached from studies on the \(\alpha \)-fetoprotein genes of the rat and the mouse; see Jagodzinski et al, Proc. Natl. Acad. Sci. USA, \(\frac{78}{28} \) (1981) 3521-3525, and Gorin et al, J. Biol. Chem. 256 (1981) 1954-1959.

The complete nucleotide sequence of human serum 20 mRNA has been determined from recombinant cDNA clones and from a primer-extended cDNA synthesis on the mRNA 2,078 nucleotides, comprises template. The sequence starting upstream of a potential ribosome binding site in the 5'-untranslated region. It contains all the 25 translated codons and extends into the poly(A) at the 3'-terminus. Part of the translated sequence codes for a hydrophobic prepeptide met-lys-trp-val-thr-phe-ile-serleu-leu-phe-leu-phe-ser-ser-ala-tyr-ser, followed by a basic propeptide arg-gly-val-phe-arg-arg. These signal 30 peptides are absent from mature serum albumin and, so far, have not been identified in their nascent state in humans. A remaining 1,755 nucleotides of the translated mRNA sequence code for 585 amino acids which are in agreement, with few exceptions, with the published amino 35 acid data for human serum albumin. The mRNA sequence verifies and refines the repeating homology in the triple-

domain structure of the serum albumin molecule.

DETAILED DESCRIPTION OF THE INVENTION-

Human serum albumin cDNA is cloned into the PstI site of plasmid pBR322 by the oligo(dG)-oligo(dC) tailing technique. Plasmid DNA was isolated from 97 positive colonies which hybridized to the enriched albumin cDNA probe, and the recombinant plasmid pHA36 was found to contain the largest insert of an albumin cDNA sequence. Its restriction endonuclease map is shown in the drawing, together with a restriction map of the primer-extended plasmid clone pHA206. The latter was obtained in a second transformation experiment after initiating the cDNA synthesis from an internal primer. This primer was a 91 base pairs long DNA fragment, MspI(152)-TaqI(182/3), isolated from pHA36. The two plasmids, pHA36 and pHA206, share 0.15 kb of homologous DNA. Together, they encode the entire sequence for human serum albumin, starting with the CTT codon for leu -10 of the prepeptide and extending into the 3'-untranslated region of poly(A).

Sequence of the Albumin cDNA. The sequence was determined for the most part on both DNA strands to ensure accuracy. All of the restriction sites used to end-label DNA fragments were sequenced across by 20 labeling a neighboring restriction site. The entire nucleotide sequence of the serum albumin mRNA, as determined from the cloned DNA in pHA36, pHA206, and from the primer-extended cDNA at the 5'-terminus of the message, is shown in the following Table 1. The inferred amino acid sequence is also indicated. The mRNA length is 2,078 nucleo-25 tides, of which 38 represent the 5'-untranslated region, 54 identify a prepeptide of 18 amino acids, 18 identify a propeptide of 6 amino acids, 1,755 code for the known 585 amino acids of serum albumin, 189 make up the 3'-untranslated region and 24 are the poly(A) sequence. Nucleotides 5 to 15 (-34 to -24) in the 5'-untranslated region (Table 30 1) are complementary to a 3'-terminal region of eukaryotic 18S RNA [Azad, A.A. and Deacon, N.J. (1980) Nucl. Acids Res. 8, 4365-4376] and thus could represent a ribosome binding site:

(5')...T T^CT C T T C T G T......albumin mRNA (3')...G A G G A A G G C G U C C m⁵_A m⁵_A......18S RNA

35

The translated portion of the mRNA sequence codes for the signal peptide and the main body of the albumin polypeptide chain. The

signal peptide is composed of a hydrophobic prepeptide of 18 amino acids and a basic propeptide of 6 amino acids (Table 1). Since prepeptides are removed from nascent secretory proteins (11ke albumin) in the endoplasmic reticulum, they are seen only in vitro in heterologous 5 translation systems. As yet, they have not been found within cells [Judah, J.D. and Quinn, P.S. (1977) FEBS 11th Mtg., Copenhagen 50, 21-29; and Strauss, A.W., Donohue, A.M., Bennett, C.D., Rodkey, J.A. and Alberts, A.W. (1977) Proc. Natl. Acad. Sci. USA 74, 1358-1362]. This is the first report of the presence and the sequence of a prepetide for human serum albumin. As it is with other secretory proteins, the conversion of proalbumin to albumin takes place in the Golgi vesicles, and the enzyme responsible for this cleavage is probably cathepsin B [Judah, J.D. and Quinn, P.S. (1978) Nature 271, 384-385]. This is also a first report on the sequence of the propeptide for normal human serum albumin.

At the 3'-end of the message, the putative polyadenylation signal sequence, AATAAA, is located 164 nucleotides downstream from the amino acid termination codon TAA and 16 nucleotides upstream from the beginning of the poly(A) sequence. Another characteristic sequence located near the polyadenylation site has been identified by Renoist, et al. [Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142]; the concensus sequence from several mRNAs was concluded as TITICACTEC. A similar sequence, TITTCCTEGT, is located 19 nucleotides upstream from the AATAAA hexanucleotide in the bumnin mRNA (Table 1).

TABLE 1

	(30)	(021)	(1560)	(350)	(440)	(330)	(620)	(710)	(300)
5	AGC		요음성	8 3 E	017 CCA		문물중	200 10,30 10,10	53 a f s
5	a E	T C	# E	A th	AAC	120 120 120 120 120 120 120 120 120 120		1ys AAG	و کی
	3 2	AAT	alu phe CAA TTT	ala thr GCA ACT	asa CAC	TAC TAC	168 169 glu cys cys GAA TGT TGC	3 5	# E
	å E	골종	ACT T	18 E	GAT		급성	ž Š	÷ 8
	-10 lle ser leu leu phe leu ATT TCC CTT CTT TTT CTC	골종	CT A	έş	Jys AAA	lys lys tyr AAA AAA TAC	ž ž	들왕	55 E
10	5 3 E	÷ 5	glu val GAA GTA	75 lys leu cys thr AAA TTA TGC ACA	S F	phe leu TTT TTG	å E	\$ ₹	₹.₹
	2 22	3 5	83n AAT	3 £	동	a E	ر ور	ala lys GCC AAA	5 55
	# E	GAT S	CTC CTC	1ys	leu TTG	まる	eg ja	36T	a E
	a E	\$.₹	ΞĚ	8 3 5 3	ag 2	충용	lys ala AAA GCT	7 5	AGA AGA
	و ۾ ه	a E	₹ ¥	£ 5	105 105 105	glu glu thr GAA GAG ACA	160 arg tyr AGG TAT	ala ser GCT TCG	gin arg phe pro lys CAG AGA TTT CCC AAA
15	-18 to p r o Het lis tro val tiu phe lle ser Het lys tro val tiu phe lle ser ATT TOC CONTINUEZONO. ATG ANG TOG GTA ACC TIT ATT TOC	10 arg ply wal pite arg arg asp ala his tys ser glu wal ala his arg pite tys asp leu oly glu glu AGG GGT GTT CGT CGA GAT CGA CAC AAGA GGT GGT CAT CAT CGT TTT AAA GAT TGG GA GAA GAA	30 phe ala gin tyr leu gin gin cys pro phe clu asp his val iys leu val asn giu val TTT GCT CAC TAT CTT CAC CAC TGT GCA TAT GAT GAA GTA	60 62 75 giu aan oya aap lya ser leu his thr leu phe qly aap lya leu oya thr val ala thr cu An Ant TGT GAC ANT CAC CTT CAT CAC ACT CTACA ACT TGT CAC ACT CACA ACT TGT CACA ACT CTACA ACT ACT	90 91 by city giv met als asp cys cys at a tys gin glu pro city are asn ciu cys nhe teu cin his tys aso asp and city city con cys at a tys gin glu pro city are asn city cys the cyt city con cyt cut cyt. The cost cyt cyt cyt cyt cyt cyt cyt cyt cyt cy	thr ala phe his asp asn ACT GCT TTT CAT GAC AAT	150 tyr phe tyr ala pro glu leu leu phe phe ala 199 arg tyr 194 ala ala phe thr 7AC TTT TAT GCC CGC GAA CTC CTT TTC TTT GCT AAA AGG TAT AAA GCT CGT TTT ACA	130 AG	220 trp ala val ala arg leu ser TGG GCA GTA GCT CGC CTG AGC
	5 5	S F	CAT S	3 E	AAT	. a 27	<u>s</u> , ₹	5 8	3 5
	AAG	e 5	8 2	¥ S	ž Ş	M S	ala lys GCT AAA	alu aly GAA GGG	5 2
	-18 Het lys ATG AAG	ral GT	글목	th S	9 9	# E	# E	GAT	5 5
20	Y.	glu val GAG GTT	# E	3 E	glu pro aly GAA CCT GGG	ala CCT	leu phe phe CTT TTC TTT	glu leu arg GAA CTT CGG	ct val
20	55	AGT SE	£ 5	15 ₹ 10 ₹	중중	kg ta	3 F	3 €	- కే
	2001	1ys	4 8 TOT	AA A	ala lys gin GCA AAA CAA	124 met cys ATG TGC	leu CTC	2.8	5 t
	8	F S	£ 95	gs CVC	1ys AAA	ATG #	pro glu CCG GAA	lys leu asp AAG CTC GAT	lys ala AAA GCA
	Ď	e S	e de CAG	69 693 1G1	1 25	crc	200	⊒ CTC	1 ys
25	3	g ₹	30 34 phe ala gin tyr leu gin gin cys TTT CCT CAG TGT	asn	2 8 E	asp CAT	÷ 55	1, s	210 phe gly glu arg ala phe TTT CGA GAA AGA GCT TTC
	CTG1	- 25	8 ₹ ¥	8 g 8	90 ela esp cys CCT CAC TGC	120 val GTT	150 phe tyr TTT TAT	180 leu leu pro CTG TTG CCA	210 ala CCT
	1701	5 5	CAG CAG	glu ser ala GAG TCA GCT	8 3	16.05	# F	3 5	ř. Ş
	Ĕ	° & E	e P	5 5	કે 🕏	£ 55	Tyr.	3 E	€ ₹
	8	val pr	ξE	5 3	aly glu met GGT GAA ATG	arg leu val arg CGA TTG GTG AGA	ala arg arg his pro GCC AGA AGA CAT CCT	177 ala ala eys GCT GCC TGC	g g
30		913	د 55	ala asp GCT GAT	중중	ra c	5 5	£ 33	# E
		9- 6- 59V	4 ie	el e	g 3	3 5	\$ 5	al a	AAA I
		- # 5	3 E	ct val	thr tyr ACC TAT	5. P.	ξŞ	<u>\$</u> ₹	leu gln CTC CAA
		1 - 6 p r o -1 1 ser als tyr ser als tyr ser arg djy val phe arg arg arg arg arg arg ard his lys ser rec got raf rec Aca Gef Gf6 TIT cef CoA CAT GCA CAC AcA Aci	21 ala leu val leu Ile ala GCC TTG GTG TTG ATT GCC	53 57 101	AC th	. E 55	- 28	190 190 als als age leu leu pro lys leu asp giu leu arp asp giu dol ty lys als ser aer als lys gin are leu lys con cor	201 als see leu gin lys phe gly glu erg als phe lys als trp als val als ang leu ser gin arg phe pro lys als olu phe als GCC AGT CTC CAA AAA TIT GGA GAA AGA GGT TTC AAA GCA TGG GCA GAA GGA GCC CGG CAG AGA TTT CCC AAA GCT TGC AAA GCT TTC AAA
25		8 1 2	3 5	草草	륁გ	3 E	141 glu 11e GAA ATT	ર કુ	AGT
35		750	21 818 GCC	2 5. \$\	91.9 CCT	111 assn AAC	£ 58	12 e 5 20	201 818 600

	(890)	(980)	(10701)	(1160)	(1250)	(1340)	(1430)	(1520)	(1610)	530 val GTT (1700)
	260 leu CTT	290 11¢ ATT	320 ala GrT	350 GCC GCC	Jeu CCT	ara CGT		\$70 AGT	1 ys	530 val GTT
5	€ 3	289 107	tyr TAT		£ 5			CT A	£ 55	2 E
	2 S	5 PS		ser väl val leu leu leu ara leu TCT GTG GTG CTG CTG AGA CTT	5 X	₹	438 cys TGT	r S	ral GT	° 55
	914 904	ser his TCC CAC	13 A A	CTC	ã E	3E)	437 lys cys AAA TGT	AG th	1 × 2	Ę t
	G SP	lys AA	316 asp val cys GAT GTT TGC.	val leu leu leu GTG CTG CTG	8. Y	واھ 90	1ys	1ys	ž Š	동종
-10	G Sp	lys pro leu leu glu lys AAA CCT CTG TTG GAA AAA	val CTT	of CTC	ala 1ys val phe asp GCC AAA GTG TTC GAT	ass AAT	aly ser lys GGC AGC AAA	£ 95	콩	lys lys AAG AAA
	ala CCT	3 5		cre cre	ag E	£ 55		ž ž	8 2	lys AAG
	253 glu cys GAA TGT	1 e	ser lys AGT AAG	val	val GTG	phe TTC	crc crc	3 E	val GTC	1 = 2 ×
		pro leu CCT CTG	Ser AGT	3e TCT	A A	1, ys	.s. ₹	val	ala leu glu val GCT CTG GAA GTC	등중
	3 F	1, s		τ <u>ζ</u> τ Σ	چ و	400 alu tyr GAG TAC	430 1eu aly CTA GGA	\$61 1GT	5 CT	ž Ş
15	250 gly asp leu leu GGA GAT CTG CTT	278 279 280 cys cys qlu lys TCC TGT GAA AAA	310 asp phe val glu GAT TTT GTT GAA	340 arg hls pro asp tyr AGG CAT CCT GAT TAC	369 370 glu cys tyr ala lys val GAA TGC TAT GCC AAA GTG	충음왕	430 ser arg asm leu dly lys val TCA AGA AAC CTA GGA AAA GTG	ago agi ser val val leu asn gin leu cys val leu his qiu lys thr TCC GTG GTC CTG AAC CAG TTA TGT GTG TTG CAT GAG AAA ACG	490 phe ser ala TTT TCA GCT	520 ser glu lys glu TCT GAG AAG GAG
	S 5	279 cys TGT	# E	5 TO	369 atu cys GAA TGC	leu aly CTT GGA	as AAC	45	7. Z	1ys AAG
	his gly CAT GGA	278 100 100	g 35	CAT CAT	~ 중 중	CT Le	ser arg asn TCA AGA AAC	leu asn CTG AAC	ξE	물왕
		278 279 leu lys qlu cys cys CTG MG GM TGC TGT	pro ser leu ala ala asp CCT TCA TTA GCT GCT GAT	ala arg arg his GCA AGA AGG CAT	pro his CCT CAT	alu aln GAG CAG	5 5	CTC	asn arg arg pro cys phe ser AAC AGG CGA CCA TGC TTT TCA	75
20	245 246 glu cys cys GAA TGC TGC	lys leu lys AAA CTG AAG	ala ccT	ala arg GCA AGA	50	£ 8	glu val GAG GTC	val GTC	2 5	510 514 phe his ala asp lie cys thr leu TTC CAT GCA GAT ATA TGC ACA CTT
20	245 Cys	3 E	ser leu ICA TTA	ele SS	360 361 cys cys ala ala ala asp TGC TGT GCC GCT GCA GAT	# E	9 8	5 CTC	\$ 5	F S
		₹ ₹	P 57	glu tyr GAA TAT	£ 5	atu teu GAG CTT	G val	leu ser CTA TCC	ACG	5, 57 7, 50 1, 50
	his thr CAC ACG	270 asn gin asp ser lle ser ser lys AAT CAA GAT TCG ATC TCC AGT AAA	50		5 ts		3 F	tyr leu TAT CTA		= \$
	54 S	- 56	300 ale asp leu GCT GAC TTG	330 phe leu tyr TTT TTG TAT	ala SS	390 392 gln asn cys CAA AAT TGT	th.	tyr T∧T	5 c c c	S10 hIs ala asp CAT GCA GAT
25	240 1ys val AAA GTC	270 ser 11e 7CG ATC	asp GAC	3 5	361 933	390 gln asn CAA AAT	£ 55	450 glu asp GAA GAC	1 teu	₹ 5
	240 1,53	270 3eF TCG	3 s s	330 TT	360	390 919 CA	420 thr	85 45 84 45	\$er 700	S10 F1s
	240 leu thr lys val CTT ACC AAA GTC	gln asp CAA GAT	asn asp glu met pro ANT GAT GAG ATG CCT	leu gly met TTG GGC ATG	giu thr thr leu glu lys GAA ACC ACT CTA GAG AAG	gin asn leu lie lys CAG AAT TTA ATC AAA	25 Z	ete cys ala TGT GCA	£ 8	af F
	3 E	- ₹ ₹	ATG	58	- 5	= =	12.5	648 Cys	¥ ¥	phe thr
	es p	glu asn GAA AAT	98	phe leu TTC TTG	3 5	3 ¥	f &	£ 5	476 477 cys cys TGC TGC	a L
30	草袋	16 8	S S		45 F	AAT	5 8	arg met AGA ATG	476 578 700	ala glu thr GCT GAA ACA
	510	265 6.ys	glu asn GAA AAT	asp val	A S	48	val	ž Ž	1ys	£3
	3 5	11e ATC	P P	asp		5 T	1ys	1ys	A S	£ 5
	ser lys leu val thr TCC AG TTA GTG ACA	ξž	291 ala glu val GCC GAA GTG	321 glu ala lys GAG GCA AAG	351 lys thr tyr AAG ACA TAT	381 val glu glu GTG GAA GAG	420 tyr thr lys lys val pro gln val ser thr pro thr leu val TAC ACC AAC AAA GTA CCC CAA GTG TCA ACT CCA ACT CTT GTA	glu ala lys GAA GCA AAA	471 480 asp ary the just by sys by the glu ser leu val GAC ARA GTC ATG CTG ACA GAN GTC TTG GTG	501 glu phe asn GAG TIT AAT
35	3e7	261 ala lys GCC AAG	£ 8	# S	351 1ys thr AAG ACA	- 8.3	AG th	- ₹ 3	\$ 42 \$ 5	a be
35	. 23 E	261 818 000	291 ala GC	32 y 6	351 178 AAG	381 val GTG	411 TAC	4 6 5	471 asp GAC	5 5 5

	6	£	2	
	. 6	188	V2)	
	360 1ys AAG	¥c	. ≨	
5	559	T &	TCT	
•	558 559 cys cys TGC TG ^C	ACAT	AATT	
	1ys AAG	CATC	ž	
	5.5	<u> </u>	¥¥¥	
	CT val	3 5	¥	
10	# E	5 5	CTGT	
	를 는	夏至	295	
	ect s	<u>و</u> ي	ð	
	불	ect s	AAAG	
	55 g 55	8 £ 8	CTCT	_
15	8.2	AGT	5110	207
	ATG ATG	. 5	Ĕ	¥
	val CTT	ed a	101	:
	ala 601	val GT	E	8
20	1 ys	E E	VICTO	
	3 5	₹ ¥	19	ž
	£ 8	¥. ¥	È	. 2
	96 86	a 4 661	¥¥	3
	¥ <u>}</u>	5 8	5	8
25	55 th 25	52 95 86	¥g.	3
	# 55 55	. 8	3	Ž
	1ys	# E	Š	1
	£ 55	567 cys TOC	9	-
	1ys AAG	휴	ž š	1
30	£ 5	9 8	<u> </u>	1
	AAA	1ys	CCA	
	rs cro	6AT 48	žCT.	,
	550 559 560 at a 150 bit by a lat the lya glu qin leu lya ala wal wet any and he ala ala phe wal glu lya ops cys lya glu leu wal lya his lya pro lya ala the lya glu qin leu lya ala wal wet any and he ala ala phe wal glu lya ops cys lya cuc ere cre cre me cox ack and one cox cre cre me cox ack and one cox act and one cre cre cre cre cre cox ack and one cox act and one cre cre cre cre cre cox ack and one cox act and one cre cre cre cre cre cox ack and one cre cre and cre	\$60 as by giu the cys phe als giu giu giu giy lys lys leu wal als als sep and als als leu giy leu ter ter as as as p asp lys giu the cys phe als giu giu giv giy lys lys leu wal als als as exp gin als als leu giy leu ter ter conc gio gio de gir wan cur cor cor ar and care to cor tra tha carecontificande (1883)	ter ter cateracetaceateartakoloamamatamaattatteateitittetttitetttitettaataacaaceaetetaamaalaaattattita (2002)	
35	550 559 also wal lys his lys pro lys ala thr lys glu qin leu lys ala wal met sap asp phe ala ala phe wal glu lys cys cys cyc circ circ and circ mic circ not con circ mic circ mic circ circ circ and circ mic circ mic circ circ circ mic circ mic circ cir	561 813 CCT	7	(2078) 20 20 20

Following are examples which illustrate procedures, including the best moder for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

Example 1 Isolation of Messenger RNA

Human liver mRNA was obtained following the procedure of Chirgwin, et al [Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) <u>Biochemistry</u> 18, 5294-5299]. Immunoprecipitation of albumin containing polysomes was performed according to Taylor and Tse [Taylor, J.M. and Tse, T.P.H. (1976) <u>J. Biol. Chem.</u> 251, 7461-7467]. <u>In vitro</u> translation of mRNA was carried out in a reticulocyte cell-free system, following the instruction of the manufacturer (New England Nuclear). The translation products were separated electrophoretically according to Laemmli [Laemmli, J.K. (1970) <u>Nature</u> 227, 15 680-685.

Example 2 Cloning Procedures

Double stranded cDNA was synthesized as described previously [Law, S., Tamaoki, T., Kreuzaler, F. and Dugaiczyk, A. (1980) Gene 10, 53-61]. It was annealed to Pstl-linearized pBR322 DNA [Rolivar, F., 20 Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crossa, J.H. and Falkow, S. (1977) Gene 2, 95-113] that had been tailed with 15 dG residues/3'-terminus [Dugaiczyk, A., Robberson, D.L. and Ullrich, A. (1980) Biochemistry 19, 5869-5873]. The annealed DNA was used to transform E. coli strain RR1, as detailed previously [Law, S., et al., 1bid.]. The albumin clones were selected using the colony hybridization method of Grunstein and Hogness [Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965], with [32P]-labeled cDNA synthesized with the immunoprecipitated polysomal mRNA as template.

As shown in Example 5, plasmids pHA36 and pHA206 were deposited in E. coli HB101 hosts. The plasmids were obtained from E. coli RR1 hosts, described in this example, and transformed into E. coli HR101 by standard procedures well known to those of ordinary skill in this art. The E. coli RR1 hosts were lysed and then centrifuged to separate the chromosomal DNA, cell DNA and plasmid DNA. The plasmid DNA, remaining in the supernatant, is precipitated with ethanol and the precipitate is resuspended in buffer, e.g., TCM (10mM Tris-HC1, pH 8.0, 10 mM CaCl2, 10 mM MgCl2). The cells for transformation are

prepared as follows: 120 ml of L-broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl) are inoculated with an 18 hour culture of HR101 NRRL B-11371 and grown to an optical density of 0.6 at 600 nm. Cells are washed in cold 100 mM NaCl and resuspended for 15 minutes in 20 ml 5 chilled 50 mM CaCl₂. Bacteria are then concentrated to one-tenth of this volume in CaCl₂ and mixed 2:1 (v:v) with annealed plasmid DNA, prepared as described above. After chilling the cell-DNA mixture for 15 minutes, it is heat shocked at 42°C for 2 minutes, then allowed to equilibrate at room temperature for ten minutes before addition of L-broth 10 times the volume of the cell-DNA suspension. Transformed cells are incubated in broth at 33°C for one hour before inoculating selective media (L-agar plus 10 ug/ml tetracycline) with 200 ul/plates

15 Example 3 Mapping of Restriction Endonuclease Sites

Restriction endonucleases were obtained from Rethesda Research Laboratories and New England Biolabs and were used according to the manufacturers' instructions. The digested DNA fragments were analyzed electrophoretically on agarose [Helling, R.B., Goodman, H.M. and Boyer, H.W. (1974) J. Virol. 14, 1235-1244] or acrylamide [Dinman, C., Fisher, M.P. and Kakefuda, T. (1972) Biochemistry 11, 1242-12507 gels.

Plates are incubated at 37°C for 48 hours to allow the growth of

DNA Sequencing

transformants.

Example 4

35

DNA fragments were dephosphorylated with bacterial alkaline phosphatase (Worthington) and labeled at the 5'-ends with polynucleotide kinase (Boehringer-Mannheim) and \(\gamma^{32} \)]ATP. Following digestion with a second restriction endonuclease and electrophoretic separation of the fragments, DNA sequence determination was done according to the procedure of Maxam and Gilbert [Maxam, A. and Gilbert, W. (1980) Methods Enzym. 65, 499-560] and the degradation products were separated electrophoretically on 0.4 mm acrylamide gels as described by Sanger and Coulson [Sanger, F. and Coulson, R. (1978) FEBS Letters 87, 107-110].

Example 5 Recombinant Plasmids pHA36 and pHA206

As disclosed in Example 2, albumin clones were selected by hybridizing to the enriched albumin cDNA probe. Plasmid pHA36 contained the largest insert of an albumin cDNA sequence. Both plasmids pHA36 and pHA206 have been deposited in a viable <u>E. coli</u> host in the

permanent collection of the Northern Regional Research Laboratory (NRRL), U.S. Department of Agriculture, Peoria, Illinois, U.S.A. Their accession numbers in this repository are as follows:

HB101(pHA36) - NRRL B-12551

- HB101(pHA206) - NRRL B-12550

5

E. coli HB101 is a known and widely available host microbe. Its NRRL accession number is NRRL B-11371.

NRRL B-12550 and NRRL B-12551 are available to the public. upon the grant of a patent. It should be understood that the availability 10 of these deposits does not constitute a license to practice the subject invention in derogation of patent rights granted with the subject instrument by governmental action.

E. coli RR1 and E. coli HB101 are known and widely available host microbes. Their NRRL accession numbers are NRRL B-12186 and NRRL 15 B-11371, respectively.

pBR322 is a well known and widely available plasmid. It can be obtained from the following host deposit by standard procedures:

NRRL B-12014 - E. coli RR1 (pBR322).

YEp6 is a well known and widely available yeast episomal plasmid.

20 It can be obtained from the following host deposit by standard procedures:

E. coli HB101 (YEp6) - NRRL B-12093.

Example 6 Assembly of the Serum Albumin Gene

Assembling the pieces together is a straighforward task of restriction enzymology. There is only one Mspl site in the overlapping
DNA sequence of the two cDNA clones. Two enzymatic steps of (i) Mspl
digestion of the two DNAs, followed by (ii) the use of ligase, an
enzyme that seals DNA fragments, will give the desired product.
Although two other undesired DNA species will also be obtained in the
course of this recombination reaction, both of them will differ substantially in size. Thus, separation and isolation of the desired DNA
species will be achieved.

The assembled DNA clone can be used to transform two types of cells:

(a) Escherichia coli

- (b) Saccharomyces cerevisiae
- (a) The vector of choice is plasmid pBR322, the same that has

been successfully used for cloning of the two fragmented pieces of the serum albumin cDNA.

In order to transform yeast with the serum albumin structural gene sequence, the DNA must be inserted into one of the s existing yeast plasmid vectors. This can be accomplished by taking advantage of the fact that several restriction endonuclease recognition sequences are absent from the cloned serum albumin DNA. Svnthetic EcoR1 DNA linkers can be ligated to the DNA fragment containing the serum albumin sequence followed by insertion (ligation) into one 10 of the yeast plasmid vectors, e.g., YEp6, at the Eco R1 cloning site. The fused chimeric plasmid can be used to transform yeast according to an established procedure [Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Proc. Natl. Acad. Sci. USA, 75, 1929]. YEp6 can be obtained from the NRRL repository, as disclosed supra.

15 Example 7 Expression of the Serum Albumin Gene

The main body of the structural gene will be transcribed by the E. coli or yeast enzymes. If little or no albumin is produced with the selected host, then an Escherichia coli promoter DNA sequence carrying an initiation codon, i.e., ATG, can be ligated at the begin-20 ning of the serum albumin structural gene. Such elements are known and available, e.g., lac promoter used for the expression of human interferon gene in E. coli [Proc. Natl. Acad. Sci. 77, 5230 (1980)]; source of promoter DNA [Proc. Natl. Acad. Sci. 76, 760 (1979)]. Also, see Nature, Vol. 281, October 18, 1979. It has already been 25 documented that such Escherichia coli promoter sequences function well in the expression of foreign genes in Escherichia coli [Mercereau-Puijalon, O., Royal, A., Cami, B., Garapin, A., Krust, A., Gannon, I. and Kourilsky, P. (1978) Nature 275, 505; and Goeddel, D.V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Grea, R., Hirose, 30 T., Kraszewski, A., Itakura, K., and Riggs, A. (1979) Natl. Acad. Sci. USA 76, 106]. For expression in yeast, see Rose, M., Casadaban, M.J. and Botstein, D. (1981) Proc. Natl. Acad. Sci. USA 78, 2460 and 4466.

Example 8 Screening of Clones Producing Albumin

Immunological methods can be used to detect small amounts of as albumin made in a bacterium. Flat disks of flexible polyvinyl are coated with the IgG fraction from an immune serum and the disks are pressed onto an agar plate so that antigen released from an in situ lysed microbial colony can bind to the fixed antibody. The plastic

disk is then incubated with the same total IgG fraction labeled with radioactive iodine so that other determinants on the bound antigen can in turn bind the iodinated antibody. Radioactive areas on the disk expose X-ray film during autoradiography and thus identify colonies producing the protein which is being screened for. Detailed protocols of this procedure have been published [Broome, S. and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA, 75, 2746]. The purification of human serum albumin can be accomplished by using procedures well known in the art. For example, procedures disclosed in a chapter by T. Peters: Purification and Properties of Serum Albumin, in: The Plasma Proteins, Putnam, Ed. Academic Press, New York, 1975, can be used.

The work described herein was all done in conformity with physical and biological containment requirements specified in the NIH Guidelines.

-12-

CLAIMS

- 1. Plasmid pHA36, having a restriction endonuclease pattern as shown in the drawing.
- Plasmid pHA206, having a restriction endonuclease pattern as shown in the drawing.
- 3. $\underline{\text{E.}}$ $\underline{\text{coli}}$ HB101 (pHA36) having the deposit accession number 10 NRRL B-12551.
 - 4. $\underline{\text{E. coli}}$ HB101 (pHA206) having the deposit accession number NRRL B-12550.
- 5. A microorganism modified to contain a nucleotide sequence coding for the amino acid sequence of human serum albumin; said nucleotide sequence is as follows:

20

25

	(36)	20 lys AAA (170)	50 ala GCA (260)	BO leu CTT (1350)	(490)	(330)	(420)	(01.7)	(300)
	¥ 36	8 17 X	8 4 2	8 3 5		1 to		5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	
5				60 62. glu ann cys asp lys ser leu his thr leu phe city asp lys leu cys thr val als thr leu Cys Ant TGT CAC ANA TCA CTT CAT ACC CTT TTT CGA CAC ANA TTA TG2 ACA GTT CG2 ACT CTT	110 asb asm pro GAC AAC CCA	3 €	169 170 cys aln TGC CAA		<u>-</u> 5
כ	leu phe leu ohe CTT TTT CTC TTT	10 ale his arg phe lys eso leu aly alu glu asn phe GCT CAT CGG TIT AM GAT TIG GGA GAA GAA AAT TTC	aly phe CAA TTT	ala thr GCA ACT	38	¥ 5	168 cys	3 5	# E
	# E	골종	4.5	3 E	8 7	2 ₹	급종	25	- 9g
	leu phe CTT TTT	2 8	5 5	į	2.3	2.3	Ęţ	£ 8	45
	ę	÷ 5	glu val thr GAA GTA ACT	£ 55	5 S	3 2	# E	2.3	× 3
10		3 2	5.5	75 asp lys leu cys thr GAC AAA TTA TGC ACA	- 55	glu thr phe leu lys GAG ACA TTT TTG AAA	lys ala ala phe thr AAA GCT GCT TTT ACA	- 2 2 2	arg phe pro lys ala glu AGA TTT CCC AAA GCT GAG
		8 5	7 2	%.₹	3 5	4 5	25	ser ala TCT GCC	2 =
		\$ 3	3 5	B ×	ag E	28	<u>8. ¥</u>	F 5	25
		# E	\$ X	25	5 8 8	28	7.5	25	£ 9K
		9 2 8	8 - X	PRE	100 101 alu cys GAA TGC	5 2 X	0 5 5	190 lys ala AAG GCT	250 AGC 48
15		. # 5	= 5	3 5	5 4	88 2	<u>s</u> §	716	35
		# 55 55	S TA	4 2	2 3	£ 5	를 눈	23	£ 8
		7 E	23	2 5	÷ 8	å E	å E	S S	2 b
		10 glu val ala his arg phe lys asp leu ely elu glu CAG GTT GCT CAT CGG TITT AAN GAT TTG GGA GAA GAA	ξE	3 E	100 101 aia iya qin qiu pro giy arq asn qiu cys ohe leu qin his iya aso GCA AAA CAA CAC CCC GGG AGA AAT GAA TGC TTC TTG CAA CAC AAA GAT	130 ata phe his aso asn alu GCT TTT CAT GAC AAT GAA	160 pro glu leu leu phe nhe ala lya aeg tyr CCG GAA CTC CTT TTC TTT CCT AAA AGG TAT	arg asp qlu oly CGG GAT GAA GGG	E Z
20		AGT	25	7 Se	골중	A th	CH E	35	# 5 #
20		× 5	¥ 25 15	¥ ¥	- 5	124 6ys	1 E	glu leu GAA CTT	5 5
		£ 5	5.9	5 5	L XX	A TG	€.8	CAT 35	e 5
		2 Y	t 6	62 Cys	# 5	520	8 55	55	1ys
		- g 2	CH C	AAT	2 5 5	CAT AS	# 8 8	1ys	
25		-6 p r o -1 1 arg gly val phe arg arg asp ala hla lys ser Acc cct ctc ccA CAT GCA CAC AAC AGT	30 phe als gin tyr leu gin gin cys pro phe giu aap his val iys leu val aan giu val thr cie. TTT GCT CAG TAT CTT CAG CAG TTT GAA CAT CAA AAA TAA GTG AAT GAA GCT CAA	50 62. glu asn cys asp lys ser leu his thr leu bhe aly CAA AAT TGT GAC AAA TCA CTT CAT ACC CTT TTT CGA	100 101 cys cys ata lys gin glu pro gly are asn glu cys phe leu gin his lys asp TGC TGT GCA AAA CAA GAA CCT GGS AGA AAT GAA TGC TTC TTG CAA CAT GAT AGA GAT	120 129 say yel giu val asp val met cys con 11G GTG AGA CCA GAG GTT GAT GTG ATG TGC	₹ ¥ ₹	180 pro lys leu CCA AAG CTC	210 ala phe 1ya ala trp ala val ala arg leu acr gin gcr TTC AAA GCA TGG GCA GCT GCC CTG AGC CAG
		F 50	16 S	# 55 CC #	£ 5	3.8	를 E	3 5	Ē Ş
		° ž E	£ 55	£ 5	# 5	£ 5	7,7	3 t	₹ 3
		val phe CTG TTT	ã E	문장	ATG Tet	P 50	88	57. 100.	5 5
		91,	- 2 5	ala asp glu ser ala GCT GAT GAG TCA GCT	glu met ala asp GAA ATG GCT GAC	15 55	£ 5	177 ala eys GCC TGC	를 E
30		40 arg	41 e	를 는	86. 9 .	3 5	₽ 5	al s	lys AA
		- 50	leu lle ala TTG ATT GCC		17.1	£ 5	arg arg his pro tyr phe AGA AGA CAI CCI TAC TIT	<u>₹</u> ¥	gin lys phe gly glu arg CAA AAA TTT GGA GAA AGA
		ala tyr ser GCT TAT TCC	25.0	53.	F S	8 8	- # 8	\$ 5	3 L
		2 5	3 2	r y	€ §	15 C	= =	5 5	AGT AG
35		-1 -6 pro -1 1 ser ala tyr ser arg gly val phe arg arg asp ala his lys ser rec ect nat rec Acc ect ere rif cet eck exit eck cet Acc Acc	21 ala leu val leu lle ala phe ala gin GCC TTG GTG TTG ATT GCC TTT GCT CAG	51 53 . Iys thr oys val AAA ACA TGT GTT	arg glu thr CGT GAA ACC	130 is pro giu val ang pro giu val ang val and coys the als phe his and and old the phe leu lys lys tyr leu and rec pro cor it of the cor cor it of the AdA COX COX GIT GIT ATT COX OC ANT GAR GAR AND ATT ITE ANA MAN INC ITA	160 160 bil e ala arg arg his pro tyr phe tyr ala pro giu ieu ieu phe ahe ala la tys arg tyr lys ala ala phe thr ciu sys cys cin Guh III e ala arg arg his pro tyr phe tyr ala pro giu ieu ieu phe ahe ala lys arg tyr lys ala ala phe thr ciu sys Guh III coc kak Ack Chi ccr 710 TIT 171 CCC CCC GAN CTC CTT TTC CTT ACK ACK ACK ACK ACK CTTT ACK CAN TGT TCC CAN	190 11 als als as plys als als cys teu teu peo lys teu aso glu teu arg aso glu opy tys als ser aer als lys oln ara teu lys 12 oct oct GAT AAA GCT GCC CTTG CCA AAG CTC CAT GAA CTT CGG GAT GAA GAG GCT TGC TCT GCC AAA CAG AGA CTC AAG	201 als ser leu gin lys phe gly glu ang als phe lys als tep als val als ang leu ser glu and phe pro lys als diu phe als diu GCC AGT CTC CAA AAA TIT GCA CAA AGA GCT TTC AAA GCA 105 GCA GTA GCT CGC CTG AGC CAG AGA TIT CCC AAA GCT GAG CTG AGC CAA

	(890)	(980)	(1070)	(1160)	(1250)	(1340)	(1430)	(1520)	(1610)	(1700)
5	260 1eu CTT	290 11 e	320 818 67.1	£ 50	Jan CCT	87 CCT	\$\$0 \$13 \$40	&70 Ser AGT	570 173 AA	530 CTT
	GAC	289 cys TCC	asn tyr AAC TAT	3 E	910 CCT	15 E	638 cys lys TGT AAA	5 Y	570	5 E
	AGG GCG	5 P		ž ž	¥ \	3 ₹	438 cys TGT	2 2	GT C	45 67
	5. S	75 75	316 eys 1ys TGC AAA	G CT	a E	CTC	437 Cys	th AGG	7,7	ᅣ
	asp asp GAT GAC	glu lys GAA AAA	316 val eys CTT TCC	<u> </u>	₽. ₹	÷ 8	K XX	8.3		aln thr CAA ACT
10	SAT CAT	2.8	2 E	val leu leu GTG CTG CTG	lys val phe asp clu phe AAA GTG TTC GAT GAA TTT	gin asn ala leu leu val CAG AAT GCG CTG TTA GTT	7 Se 7	문용	문장	\$ \$
	2 5	leu leu CTG TTG	GAT	55	를	5.5	5 5		SS CAT	1ys
	253 glu eys GAA TGT	3 E	1ys	val	510	ag E	510	125	5. CTC	# 15 A
		280 glu lys pro leu leu GAA AAA CCT CTG TTG	AGT	10 F	1 ys	600 glu tyr lys phe GAG TAC:AAA TTC	₹ ₹	55	중중	53
15	S E	₹ ₹	골중	1 ty	÷ 8	7,5	qly lys ccA AAA	461 0ys TGT	3 5	1 2
15	250 83p leu leu GAT CTG CTT	8 3 3	5 to F2	340 85p GAT	37 14 14	95 P	5 5 T	3 E 56	490 818 GCT	910 GAG
	55 P	278 279 cys eys TCC TGT	a E	340 phe leu tyr glu tyr ala arg arg his pro asp tyr ser val val lev lev leu ara lev TIT TTG IAT GAA TAT GCA AGA AGG CAT CCT GAT TAC TCT GTC GTG CTG CTG CTG CTG AGA CTT	369 370 qlu cys tyr ala lys val GAA TGC TAT GCC AAA GTG	5.8	830 ser arg asn leu qly lys val TCA AGA AAC CTA GGA AAA GTG	58	៖ ៦	510 514 520 his ale asp tie oys thr leu ser div lys giu are din lie iys lys din thr ale leu CAT GCA CAT ATA TGC ACA CTT TGT GGC ANG GAG CAA ATC AAC AAA CAA ACA CTT GCA CTT
	his gly	278 Cy3	85 Z	arg his AGG CAT	균성	phe alu aln leu TTT GAG CAG CTT	5 5	450 glu asp tyr leu ser val val leu asn GAA GAC TAT CTA TCC GTG GTC CTG AAC	# E	5 5
	25	₹.3	al a	PTG AGG	pro his CCT CAT	투왕		55	480 ser leu val asn arg arg pro cys TCC TTG GTG AAC AGG CGA CCA TGC	3e TCT
20	245 246 895 893 TGC TGC	1ys	ala GCT	AGA AGA	55	문용	val GTC	val GTC	£ 5	E CT
	245 246 his thr glu cys cys CAC ACG GAA TGC TGC	ser lys leu lys AGT AAA CTG AAG	3 5	- 3	360 361 cys cys ala ala ala asp TCC TGT CCC CCT CCA GAT		918	val GrG	5	514 11e cys thr 1eu ATA TGC ACA CTT
	8.8	1,ys	pro ser CCT TCA	glu tyr GAA TAT	ala ala GCT GCA	3 E	val GTA	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	8 79 9 79	5 5 5 T
	후	AGF	CCT	₹ ₹	5 5	골	3 E	5 FS	asn AAC	7 t
	± 5	15	110	tyr TAT	# 55 F	392 cys TGT	pro thr CCA ACT	7.₹	val GTC	\$ P
25	55	270 ser 11c TCG ATC	300 ala asp leu GCT GAC TTG	330 met phe leu ATG TTT TTG	361 167 161	390 392 gin asn cys giu leu CAA AAT TGT GAG CTT	£ §	5 S	480 ser leu TCC TTG	510 his ale asp CAT GCA GAT
	240 AA		8 2 5	33 F	36 873 757		42 th		84 PD T	5 5 5
	Acc th	glu asn gln asp GAA AAT CAA CAT	ere cct	phe leu gly met TTC TTG GGC ATG	thr leu glu lys a	gin asn leu lle lys CAG AAT TTA ATC AAA	1 2	448 oys ala TGT GCA	중중	
	asp leu GAT CTT	58	asp glu met GAT GAG ATG	91,	gle	asn leu lle AAT TTA ATC	crc crc	448 033 TGT	476 477 cys cys thr TGC TGC ACA	A th
30	GA TA	AAT	95	3 2	T T	3 T.	£ ₹	5 2	477 cys TGC	ag 2
	leu val thr asp TTA GTG ACA GAT	2 A	asp CAT			AAT	. F S	Met ATG		₽ \$
	25 070	263 633 1GT	AAT AAT	asp val GAT GTC	glu thr GAA ACC	£ 55	CTA C	5 5 5	thr 1ys Acc AAA	€ ₹
	a E	1 e	val glu GTG GAA	SAT CAT		era CCT	£ \$	₹ ₹		at a
	1ys	ty TAT	val CTC	¥ 5.	74	26.5	420 byr thr 1y3 lys val pro gin val ser thr pro thr leu val giu val TAC ACC AAG AAA GTA CCC CAA GTG TCA ACT CCT GTA GAG GTG	ala lys arg GCA AAA AGA	520	AAT
35	5 55	261 ala 1ys GCC AAG	3 YS	321 glu ala GAG GCA	호 호	9 45 G¥ 35	th Acc	2.5	\$ 5	ag E
	£ 15 F	261 818 GCC	291 818 GCC	321 91u GAG	351 1ys	381 val GTG	41 171 170	4 e 100	471 GAC	501 glu phe asn GAG TIT AAT

	giu Leu val lya hia lya pro lya ala thr lya giu gin leu lya ala val met asp aap ohe ala ala phe val giu lya cya cya lya Gac citc gitc MA Gac MG GCC MG GCA ACA MA Gac CAA, CIG AMA GCT GIT AIG GAT GAT CAT TIT GIA GAC AAG TGC TGC. MG (1790)	S61 S80 And the cys pike also glu div diy lys lys leu val also also ser gin also also leu diy leu ter ter cor ove one one of or over over one of over over over over over over over over	ter ter Chictorochaeachamaramangamantamantamachtaitantithtithtithtitithtigestagarachtertamamantamathtithta (2003)
260	YS AG	. 3	¥
558 559 560	5 5	ž Š	Ę
558	5 S	Š	TAY.
-	4 y	CATC	2
	5 3	<u>‡</u> ₹	¥
	7 ž	2 €	Ž.
	ŘΕ	5 8	757
	i i	3 ₹	Š
	÷ 5	2 23	₹)
	a E	<u>چ</u> کے	¥
550	5 5	88 € 45 €	STG
	\$ \$	AGT SE	, F
	A G	్లి క్ర	Ĕ
	3 E	£ 5	Ē
	÷ 5	5 75	Ė
	₹ ₹	3 E	ĬĊ
	3 E	\$ ₹	Ę
	ŧš	\$ ₹	. U
	급성	45	¥
	<u>*</u>	9	SATO.
\$	ķÝ	5 4 3	2
	÷ 5	# 85 18	¥
		ã E	¥
	2 22	75. 25.	20
	¥ 5	A C	ter ter TGACAATAA
	žš	38	<u> </u>
	ž <u>š</u>	Ag	YCCY
	2 5	.g ₹	200
	giu Leu val Iya kia Iya pro Iya ala thr Iya giu gin leu Iya ala val met aap aap obe ala ala phe val giu Iya cya cya Iya GAC CTC GTG AAA CAC AAC CCC AAG ACA AAA GAC CAA, CTG AAA GCT GTT ATG GAT CAT TTT CAT CTT GTT GTA GAG AAG TGC TGT AAG	8 3 2	ETCA.
53.	물용	561 ala CCT	CATC

TCATTTTGCCTCTTTTCTCTGTGCTTCAATTAAAAAATGGAAAGAATCTAA..... 20AA (2078)

6. Nucleotide sequence of the cDNA of human serum albumin, said nucleotide sequence is as follows:

5	(170)	(260)	(350)	(440)	(330)	(629)	(710)	230 91u GAA (300)
	2 × ¥	8 శ్రీ		5 5 5	140 147	무 등 중		Syl Syl
	Phe TTC	를 는	thr	AAC AAC	2 ₹	69 107 107	200 1ys cys AAG TGT	8 Y
	AAT	골중	60 62 75 glu san cys asp lys ser leu his thr leu phe qly asp lys leu cys thr val als thr leu GAA AAT TOT GAC AAA TCA CTT CAT ACC CTT TTT CGA CAC AAA TTA TOZ ACA GTT GOA ACT CTT	100 ini oys oys ale lys gin giu pro giy ara aan diu oys bhe leu din his lys ase ase aan TGC IGT GCA AAA CAA GAA GCT GGG AGA AAT GAA TOC TTG TTG CAA CAC AAA GAT GAC AAC	ž ž	168 cys TGT	3 5	a E
10	1 sep ala his iya ser giu wal ala hisa nen phe iya aso leu diy qiu giu oxi goa coc occ and car car cot cot cot TTT AM OX TTG GOA CAN GAA	40 pro phe glu asp his val lys leu val asn glu val thr CCA TTT GAA GAT CAT GTA AAA TTA GTG AAT GAA GTA ACT	thr val	SAT CAT	120 met cys thr als phe his aso aan diu diu the phe leu lys lys tyr ArG TCC ACT CCT TIT CAT CAC AAT CAA CAC ATT TTG AAA AAA TAC	8.8	5 A	2 8
••	골중	5 × 5	ķ ţ	AA A	AA AA	ş ţ	5 8	55 CC 59
	€ 8	₽.₹	75 lys leu cys AAA TTA TGC	leu gin his lys asb TTG CAA CAC AAA GAT	1 E	# E	₹ ₹	220 ale vel ale arg leu ser pin arg phe pro lys ale QCA GTA GCT CGC CTG AGC GTTT CCC AAA GCT
	160 170	AAT AAT	T Fe	등중	å E	ala GCT	818 000	£ 5
	88 G¥	575	1ys	15	ž Š	SCT DO	10 20	를 E
15	10 arg phe 1ys CGG TTT AAA	a E	60 62 giu asn cys asp lys ser leu his thr leu phe gly asp GAA AAT TGT GAC AAA TCA CTT CTT CCA GAC	100 101 alu cys phe GAA TGC TTC	130 asn qlu qlu AAT GAA GAG	AA A	700	P 454
	를 E	1ys AAA	5 5	55 55 50 50 50 50 50	급종	160 lys arg tyr AAA AGG TAT	190 gly lys ala GGG AAG GCT	£ 35
	5 £ 8	5 to 5	e a E	등 등 중	130 AAT	6 F 8	190 178 AAG	220 AGC
	glu val ala his GAG GTT GCT CAT	CA TA	3 E	ala iya gin giu pro qiy arq aan GCA AAA CAA GAA GCT GGG AGA AAT	phe his asp TTT CAT GAC	₹.₹	5 8 8	5 5
20	£ 55	\$ 72	ž Š	£ 5	₹ ₹	a F	asp alu GAT GAA	ត្ត មួ
20	. E		£ 75	95	# E	. 4 E	გ	SG \$2
	53	a E	leu CT	oc T	ele CCT	ag E	979	6 4 5
	lys ser AAG AGT	- £ 5	ž Ž	5 ₹	ŧ 5	₹ 5	asp glu leu GAT GAA CTT	శ్లీ ప్ర
	1,ys	4 67 TGT	1ys	5 ₹	124 973 TGC	G E		t 5
25	£ 33	30 34 ale phe ala gin tyr leu gin gin cys GCC TTT GCT CAG TAT CTT CAG CAG TGT	as CVC	₹. *	124 val met cys GTG ATG TGC	150 pro tyr phe tyr ala pro glu leu leu CCT TAC TTT TAT GCC CGG GAA CTC CTT	88 CAT	210 ala phe lys ala trp GCT TTC AAA GCA TGG
	: 5	d b	62 673 101	2 5	575	50	55 CT	¥ ₹
	- g 2	3 5	AAT	90 91 cys cys TGC TGT	120 pro glu val 88p CCA GAG GTT GAT	ele SS	180 pro lys leu CCA AAG CTC	를 다
		₹ ¥			120 ca CT	5 ? F	8 5 X	210 ala GCT
		58	val ala asp glu ser ala GTT GCT GAT GAG TCA GCT	gly glu met ala asp GGT GAA ATG GCT GAC	5.5	# F	leu leu CTG TTG	5. A
30		មិន	5 5	5 P	5	7,₹	35	gin lys phe gly glu CAA AAA TTT GGA GAA
		ξE	96	ATG	arg leu val arg CGA TTG GTG AGA	5 75	177 ala oys GCC TGC	કે કે
		£ 55	SAT CAT	₹ §	val GTG	arg his AGA CAT	" 55	# F
		leu 11e TTG ATT		gly GCT	3 5	5 5	5 to	₹ ₹
35		3 5		tyr TAT	€ 5	5 5	AA A	ર્ફ કુ
33		21 ala leu val 1eu lie ala phe ala gin tyr leu gin gin cya pro phe giu asp his val 1ya leu val aan giu val thr eiu phe CCC TTG GTG TTG ATT CCC TTG CT CAG TAT CTT CAG CAG TCG CCA TTT GAA CAT CAT GAA TAA GTG AAT GAA GTA ACT CAA TTT		81 arg glu thr tyr gly glu met ala asp cor GAA ACC TAT GOT GAA ATG GCT GAC	120 aan leu pro arg leu val arg pro glu val asp val AAC CTC CCC CCA TTG GTG AGA CCA GAG GTT GAT GTG	140 glu lie ala arg ang his pro tyr phe tyr ala pro glu leu leu phe phe ala lys ang tyr lys ala ala phe thr glu oys cys GAN ATT GCC AGA AGA CAT CAT TAT IXT GCC CCG GAN CTC CTT TTC CTT GCT AAA AGG TAT AAA GCT GCT TTT ACA GAN TGT TGC	179 als als asp lys als also oys leu leu pro lys leu asp glu leu arg asp clu gly lys als ser ser als lys cln arc leu lys ccr ccr car Aan ccr ccc roc cro ric ccn Anc crc car can car cco car can can coc Anc Crr roc ann cho and crc Anc	201 ala ser leu gin 1ys phe giy giu arg ala phe 1ys ala tep ala val ala arg leu ser pin arg phe pro 1ya ala giu phe ala giu OCC AGT CTC CMA MAT ITT CGA GMA MAX GCT TTC AMA GCA GTG GCA GTA GCT CGC CTG AGC CMA TTT CCC AMA GCT CAG TTT GCA GMA
		3 5	51 lys thr AAA ACA	류종	3 5	A 16	eg s	Ser AGT
		24 818 500	2 X X	8 gr 6	111 AAC	5 g 4	£ 5 5	25 8 8 99

5	(1890)	(980)	(1070)	(1160)	(1250)	(1340)	(1430)	(1520)	(1610)	(1700)
	260 CTT	289 290 cys 11e TCC ATT	320	350 818 GCC	Jan CCT	8 to	440 113	\$70 AGT	_	530 cr 1
	aso and ala aso GAC AGG GCG GAC	289 cys TGC	asn tyr AAC TAT	3 E	ere CCT	. 15 E	₹ ¥	val GTA	Sno Dre 1ys	3 5
	£ 8	55 S	AAC AAC	5 A	1ys AAA	3 ₹	538 638 101		44	thr ala leu ACT GCA CTT
10	asb arg GAC AGG	. F 5	316 cys lys TGC AAA	. <u>3</u> E	ag E	3 5	437 638 1GT	thr pro ACG CCA	7,5	45
	\$ 5	.y. A.	¥ 8.55	3 5	asp glu GAT GAA	55 55	1ys			£ 3
	ala asp GCT GAT	pro leu leu glu lys ser CCT CTG TTG GAA AAA TCC	316 glu ser lys asb val cys lys GAA AGT AAG GAT GTT TGC AAA	val leu leu leu GTG CTG CTG CTG	ala lys val phe asp GCC AAA GTG TTC GAT	asn ala AAT GCG	, P 20	alu lys GAG AAA	glu val asp glu thr GAA GTC GAT GAA ACA	lys oln AAA CAA
	# ts	3 5		va1 GTG	A C	£ 95	충용		8 P	8, 3s
	253 glu cys GAA TGT	3 5	17s	val	val 676	축Ě	5 5	3 E	4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	Ile lys ATC AAG
15	33	£ 55	ser AGT	10	ala lys val phe GCC AAA GTG TTC	2,4	lys val AAA GTG	val leu GTG TTG	96 649	# X
	250 253 leu leu glu cys CTG CTT GAA TGT	280 qlu lys GAA AAA	륁	340 asp tyr GAT TAC	, 1 ₈	400 giu tyr lys phe gin asn ala leu leu val GAG TAC AAA TTC CAG AAT CCG CTG TTA GTT	충성	\$61 6ys		23
		270 270 ser lys leu lys glu cys cys glu lys pro leu leu glu lys ser his roc ait fic Adf AAA CTG AAG GAA TGC TGT GAA AAA CCT CTG GAA AAA TGC CAC	310 ala ala aso phe val glu ser lys GCT GCT GTT TTT GTA AGT AAG	340 phe leu tyr glu tyr ala arg arg his pro asp tyr ser val val leu leu leu aru leu TTT TTG TMT GAM TAGA AGA CAT CCT GAT TAG TCT GTC GTG GTG CTG CTG AGA CTT	370 147 147	15 PE	830 glu vel ser arg asn leu qly lys vel qly ser GAG GTC TCA AGA AAC CTA GGA AAA GTG GGC AGC	1eu	ala leu GCT CTG	520 Iya giu ara oin ile iya iya oin thr ala leu AG GAG AGA CAA ACT CAG CTT
	gly asp GGA GAT	278 279 0ys 0ys TGC TGT	ž E	Pro CCT	369 alu cys GAA TGC	390 392 gin asn cys glu leu phe alu ain leu gly CAA AAT TGT GAG CTT TTT GAG CAG CTT GGA	asn AAC	4 8	P 2	5 S
20		278 lys glu cys AAG GAA TGC	300 ala asp leu pro ser leu ala ala asp GCT GAC TTG CCT TCA TTA GCT GCT GAT	arg his AGC CAT	종	e CT	arg asn AGA AAC	AAC AS	å E	
	246 cys h1s TGC CAT	£ 3	- 2 5	A SC	S F	동웅	36 77	De CTC	165	5 E
		. ¥.	£ 5	AGA AGA	5 73	5 8	va1 CTC	c CTC	25	lou ser CTT TCT
	245 his thr glu cys CAC ACG GAA TGC	lys leu AAA CTG	300 ala asp leu pro ser leu GCT GAC TTG CCT TCA TTA	glu tyr ala GAA TAT GCA	360 361 cys cys ala ala ala asp pro TGC TGT CCC CCT GCA CAT CCT	leu phe alu CTT TTT GAG	val glu GTA GAG	tyr leu ser val val leu aan TAT CTA TCC GTG GTC CTG AAC	£ 5	¥ Ş
	his thr giv	£ ₹	5 5	ξ¥	ද ફ	3 5	val GTA	÷ 5	AGC ATG	5 5 5 7 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
25	£ ŏ	AGT	50	5 €	85 PS	5 5	3 €	leu ser CTA TCC	AAC AAC	514 11e cys ATA TCC
	₹ 5	5 5	11G	17.	ال 93	392 cys TGT	rt th	7.4	2 S S	8 35 F 15
	240 lys val AAA GTC	270 ser 11e TCG ATC	8 3 24	330 phe leu tyr TTT TTG TAT	361 cys TGT	390 392 gln asn cys CAA AAT TGT	£ 5	SA CA	3 5	- 5
	240 AX						thr ACT	\$50 91c GAA	676 477 460 thr lys eys thr glu ser leu val asn arg arg pro eys phe ser ACC AAA TGE TGE ACA GAA TCE TTG GTG AAC AGG CGA CCA TGE TTT TCA	510 glu thr phe thr phe his als asp lle cys thr lou GAA ACA TTC ACA GCA GATA TCC ACA CTT
	leu thr CTT ACC	265 cys glu asn gln asp TGT GM AAT CM GAT	met pro ATG CCT	phe leu gly met TTC TTG GGC ATG	glu lys CAG AAG	gin asn leu ile lys CAG AAT TTA ATC AAA	10 Ser	2 Y	5 €	\$ E
30	asp leu GAT CTT	48	glu met GAG ATG	leu gly TTG GGC	중중	asn leu 11e AAT TTA ATC	ra1	948 033 TGT	¥ Ş	thr phe ACC TTC
	val thr asp GTC ACA CAT	asn AAT	asp glu GAT GAG	1 e	thr leu ACT CTA	₹ E	gin val CAA GTG	50	673 TGC	T C T
	草支	₹3			ACT SP	asn AAT	. 50	ATG AT	476 477 cys cys TGC TGC	5 5
	510		glu asn GAA AAT	val GTC	AC th	£ 95	val pro GTA CCC	arg met AGA ATG	ž.≱	glu thr GAA ACA
35	3 €	1 e	95	asp CAT	8,8	£ 53	lys AAA	17s	th.	5 to
	ser lys leu TCC AAG TTA	261 ala lys tyr lle GCC AAG TAT ATC	55	133	7.5	58	thr lys lys ACC AAG AAA	glu ala lys GAA GCA AAA	val t	asn a
	75	lys AAG	5 ₹	818 CCA	r y	급성	Ę Ş	918	25	phe a
	2 2 2	261 818 600	291 818 GCC	321 glu GAG	351 173 AAG	381 val glu GTG GAA	17,7	£ 500	471 889 arg GAC AGA	501 glu phe asn ala GAG TTT AAT GCT

5	11790)	540 Ser 770 ter Is aspess bys giu thr cys phe ala giu giu qiy lya bys leu val ala ala are gin ala ala leu qiy leu ter ter Cr GuC GAT ANG GAC ACT CTT CCC GAC GGG GGT ANA ANA CTT GTT GCT CCA AGT CAN GCT CCC TTT GGC TTTA TAA CATCACATTTAAAAG (1883)	ter ter Ateracciacataraharaharahanantanantahakettattattaticittiticitticettöstgimmöcamorccigicimmamacatamittettym (2002)
	_	9¥	¥
	558 559 560 cys cys 1ys TCC TCC AAG	T te	TE .
	538	ACAT	WAT
10	1ys AAG	CATC	Ç
10	56	± ₹	¥
	CT A	2 €	7 <u>7</u>
	₹ 11 Pe	출성	500
	9 5	E E	VCAC
15	et s	5 00	32
	를 E	eg to	¥
	55	8 5 3	cctc
	8 ₹ 7	ser AGT	1100
	ATG	£ 5	E
20	4	÷ 5	110
	818 CCT	3 E	1115
	₹.	3 5	5
	3 5	2.3	¥
25	£3	2.3	5
	48	£8	₹
	2.5	- 53	CATC
	7 th 76	2 4 2	Ę.
	# 8	- 58	GAA.
30	1. X	£E.	Ž.
	20	2 2 2	, Š
	¥ 4	£ 5	SAT C
	£ 3	48	£ £
35	1,1	¥ 5	TACC
55	va1	8 8	P
	1 5	8.9	55
		5 - 5	- 5

TCATTTTGCCTCTTTTCTCTGTGCTTCAATAAAAAATGGAAAGAATCTAA..... 20AA (2078)

 Nucleotide sequence coding for the prepeptide of human serum albumin, said nucleotide sequence is as follows:

		(00)	
		-18 p r o -10 Het lys trp val tlu phe Ile ser leu leu beu phe leu ohe ser OCTITICICITCICICAMCCCCAMACCCTITGZAMA ATG ANG TGG GTA ACC TIT ATT TCC CTT TIT CTT TIT ACC	
		å E	
		38	
		# E	
		ួ≩5	
		6 3 F	
		# 5	
		# F	
		# E	
		° = 5	
		- 5 E	
		100	
		* * *	
		A # 55	
		Sec	
		100	
		55	
		NO.	
		Ş	
		425	
		101	7 \$ 8
		1010	, e
		1113	val phe GTG TTT
		8	2 5 5
			5.5
			6. 50 S
			er ale tyr ser arg gly val phe arg ard TCG GCT TAT TCC AGG GGT GTTT CGT CGA
			₹
			- 8
			ğ 5

8. Nucleotide sequence coding for pro human serum albumin, said nucleotide sequence is as follows:

5	(170)	12601	(350)	(990)	(334)	(624)	(710)	230 phe ala qlu TTT GCA GAA (300)
	8 1 X	Տ류정	8 3 E	5 5 5 5 5 5	1 to 1 to 2	동특절	200 cys	8 al 23
	T Se	ag E	ACT TA	AAC	골	168 169 cys cys TGT TGC	AAG AAG	2 5
	AAT	glu phe CAA TTT	80 818 thr leu GCA ACT CTT	88 5	7,7	168 cy3 TGT	3 €	a E
	als. SA	ACT TA		SS CAT	1ys	₹ §	2 5	4 S
10	골종	glu val GAA GTA	y th	AA A	1 ys	ž Ž	£ 95	
	충성	급정	₹ % 5	55 S	3 2	· #E	₹ ₹	Lys AA
	3 5	AAT	3 €	듬중	4 F	313	* D	£ 55
	\$ 2	510	75 1ys leu ays AAA TTA TGC	3 5	ž Š	SC \$	15 F	ᇂE
15	AA A	2 £	\$ 5 5		98	lys ala ala phe AAA GCT GCT TTT	F 5	gin arg the pro lys ala CAG AGA TTT CCC AAA GCT
	# E	NA AA	5 5	101 105 101	골종	160 arg tyr AGG TAT	ala GCT	58
	10 ala his arg phe lys aso leu aly diu glu asn phe GCT CAT CGG TIT AAA GAT TTG GGA GAA GAA TTG	40 val lys leu val asn GTA AAA TTA GTG AAT	RAE	110 glu cys phe leu gln his lys asp asp asn pro GAI TGC TTC CAA CAC AAA GAT GAC AAC CCA	130 AAT	25 P 50	130 AC 130	250 VGC 364
	₹ 5	asp his GAT CAT	말	AT TA	88	₹.¥	alu aly GAA GGG	3 E
	5 to	asp GAT	th ACC	5 5 5	CA TA	÷ 5	골종	37.9
20	2 F2	콯죓	£ 5	5 8	# E	者E	g S	als SC 13
	-1 1 arg asp ele hís lys ser glu val CGA GAT GCA CAC AAG AGT GAG GTT	an gin gin oya pro phe ciu asp his val lys leu val asn CAG CAG TGT CCA TTT GAA GAT GTG AAT	60 62 82 ale as giu as 195 ser leu his thr leu phe giy asp 195 leu cys thr vel ct fat CAG TCA CAG AAA TG GC AAA TGA CTT CAT CAG CTT TTT GCA CAC AAA TAT TGC CACA AAA CTT CACA CTT TTT CCA CACAAAA CACAAA TAT TGC CACAAAAA CACAAAAAAAAAA	90 91 oys cys ala lys gin giu pro gly aro asn TGC TGT GCA AAA GAA GAA CCT GGG AGA AAT	120 glu val asp val met cys thr als phe his asp asn glu glu thr phe leu lys lys tyr cyc cri cyr cir Ari rcc Aci cor TIT CAI CAC ANI CAA CAG ACA TIT ITG AAA AAA TAC	150 arg arg his pro tyr phe tyr ale pro glu leu leu phe phe ale lys arg tyr lys als ale phe thr clu AcA ACA CAT CCT TAC TIT TAT CCC CCG CCA CTC CTT TIC TIT CTT AAA AGG TAT AAA GG GCT TIT ACA GAA	190 Its ais ais eys feu feu pro lys feu asp glu feu arg asp dlu gly lys ais ser ser ais lys oln ara feu lys AAA gcT GCC TGC CTG TTG CCA AAG CTC GAT GAA CTT CGG GAT GAG GGT TGG TCG TGG AAA CAG CTG AAG	210 glu arg ala phe lys ala trp ala val ala arg leu ser can ack GCT TTC ANA GCA TGG GCA GTA GCT CGC CTG AGG
	AGT AGT	- £ 5	ئے <u>ہ</u>	쿰췽	ACT TA	3 €	3 €	÷ 5
	1ys	¥ \$ 5		둫중	124 val met cys GTG ATG TGC	pro glu leu CCG GAA CTC	ole AA	135
25	÷ 5	£ 82	\$ 35 5	1, s	ATG	₹,5	SAT AS	£ 5
23	- 8	48	62 1GT	2 8	val	£ 99	lys leu AAG CTC	¥ ¥
	-1 1 arg asp ele his CGA GAT GCA CAC	3 E	asn AAT	92 101	120 val asp GTT GAT	÷ 58	178	af 5
	- 58	30 gin tyr leu gin gin gys CAG TAT CTT CAG CAG TGT	용물성	90 97 7GC	120 val GT	150 phe tyr ala TTT TAT GCC	8 t 2	210 a1a GCT
	F 5	£ 8	ser ala TCA GCT	GAC CAC	pro glu cca cac	를 E	leu leu CTG TTG	5 Q
30	pro glyval phe GCT GTG TTT	e e cc	₽ 5 2	glu met ala GAA ATG GCT	£ 55	his pro tyr CAT CCT TAC	- CT	중중
	2.50	Phe TT	asp glu GAT GAG	met ATG	A 25	5 75	F %5	£ 8
		دا و	es CAT	gly glu met ala asp GGT GAA ATG GCT GAC	arg leu val arg CGA TTG GTG AGA	÷ 5	- 50	# E
	9 E S	11e		9. y	3 E	5 5 5	ala GCT	gin lys CAA AAA
		val leu fle ala phe ala GTG TTG ATT GCC TTT GCT		7.7	pro arg leu val arg pro CCC CGA TTG GTG AGA CCA	۶. و م	2, ₹	동장
35		30 31a leu val leu lle ala phe ala gln tyr leu GCC TTG GTG TTG ATT GCC TTT GCT CAG TAT CTT		th Ac	F 25	# 8	GAT A	201 als ser leu gin lys phe gly glu amp sla phe lys als tep als val als ang leu ser gln ang phe pro lys als qlu phe als qlu coc AGT CTC CAA AAA TIT GOA GAA AGA GET TTC AAA GCA TGA GCA GTA GCT CCC CTG AGC CAGA TIT CCC AAA GCT GAG CAG CAG CAG CAG CAG CAG CAG CAG CAG
		15 T	51 Iys thr AAA ACA	81 arg glu CGT GAA	111 asn leu AAC CTC	141 glu 11e GA ATT	8 to	Ser AGT
		2 to 50	2 ₹ \$	8 grg	11 AAC	₹ § §	171 818 GCT	2 4 8

5	(890)	(980)	(1070)	(1160)	r125n1	(1340)	(1430)	(1520)	(1610)	(1700)
	260 CT 162	290 11e ATT	320 818 07.T	350 ala GCC	380 CCT	410 ara		\$70 ACT		530 611
	ala asp OCC CAC	289 105	1 X		£ 55				500 pro 1ys CCC AAA	. 5 1eu v
		£ 5	asa AAC	Ĕ, Q	₹ ₹	P T	438 cys TGT	Dro val CCA GTA		als leu GCA CTT
10	F 5	10.0	316 cys lys TGC AAA	1 e	# E	2 C C C C C C C C C C C C C C C C C C C	437 438 cys cys lys TGT TGT AAA		17c	thr a
	253 glu cys ala asp asp arg GAA TGT GCT CAT GAC AGG	lys AAA	3.5 5.55 7.55	15 CTC	₹.8	# £		¥ 3×	thr t	CA A
	\$ 2	골중	2 E	leu CTG	£ 5	₹ ¥	19 S	5 5	SA A	5. ₹
	5 5	1eu	CAT	575	phe asp TTC CAT	듬왕	÷ 20	\$ 5	88 P	7 × ×
	253 573 TGT	1 E	1ys	** CTC	val 676	를 E	55	3 E	18 Z	1 2 2
15		lys pro leu leu glu lys ser AAA CCT CTG TTG GAA AAA TCC	glu ser lys asp val GAA AGT AAG GAT GTT	ser val val leu leu leu arg leu TCT GTC GTG CTG CTG AGA CTT	lys val phe AAA GTG TTC	lys phe AAA TTC	lys val qly AAA GTG GGC	55	48	5.3
	gly asp leu leu GCA CAT CTC CTT	278 279 280 oys cys qlu lys pro leu leu glu lys ser TCC TCT CAA AAA CCT CTG TTG CAA AAA TCC	윰	1 ty	* 55		436 asn leu qly lys val qly ser lys AAC CTA GGA AAA GTG GGC AGC AAA	10,3	480 glu ser leu val asn arg arg pro cys phe ser als leu glu val asp dlu thr tyr val CAA TCC TTG GTG AMC AGG CGA CGA TGC TTT TGA GGT CTG GAA GTG GAT GAA AGA TAG GTT	520 glu ara ain 11e iya iya ain thr GAG AGA CAA ATC AAG AAA CAA ACT
	250 asp leu CAT CTG	280 41 CA	310 phe val TTT CTT	35 55 57	37 1,47 1,47	400 leu aly glu tyr CTT GGA GAG TAC	436 Ieu aly CTA GGA	1 eu	818 CCT	25 25 25 25 25 25 25 25 25 25 25 25 25 2
	\$ 2	278 279 glu oys cys GAA TGC TGT	ala asp phe GCT CAT TTT	£ 55		5 8	asn AAC	£ 55	15 E	86.5
20	÷ 5	278 275 700	8 P	K 15	급정	3 5	\$ 5 \$		# E	glu lys GAG AAG
20	245 246 glu cys cys his GAA TGC TGC CAT		310 81a 8sp leu pro ser leu 81a 81a 8sp GCT GAC TTG CCT TCA TTA GCT GAT	330 phe leu tyr glu tyr ala arg arg TTT TTG TAT GAA TAT GCA AGG AGG	pro his glu CCT CAT GAA	5.5	glu val ser arg GAG GTC TCA AGA	val val leu asn GrG GTC CTG AAC	\$ 55	510 phe his ala asp lie cys thr leu ser glu lys TTC CAT GCA GAT ATA TGC ACA CTT TGT GAG AAG
	2 5 2 5 2 5 3 5	lys leu lys AAA CTG AAG	pro ser leu ala CCT TCA TTA GCT	E S		5 5	glu val GAG GTC	CTC CTC	2 5	2 E
	245 91u eys GAA TGC	lys leu AAA CTG	₹.	- ; 5	ala asp GCA CAT	leu phe CTT TTT	gle GAC	cro	£ 8	£ 5
			ž 5	glu tyr GAA TAT	360 361 cys cys ala ala ala asp TGC TGT GCC GCT GCA GAT	3 E	5 th	36 707	819 AGC	514 esp lle cys thr CAT ATA TCC ACA
25	hls thr CAC ACG	AGT A	£ 5	8 8	8.13 CCT	5.3	3 5	CT te	AAC AAC	- 5
	ž 5	7 2	3 5	žξ	£ 8	392 161	pro thr CCA ACT	asp tyr leu GAC TAT CTA	5 STC	5 S
- ·	240 lys val AAA GTC	270 ser 11e TCC ATC	300 ala asp cct GAC	330 phe leu tyr TTT TTC TAT	26. 10.	390 392 gln asn cys CAA AAT TGT	£ 5	\$ 5	3 E	# YS
	240 178	270 asp_ser 11e GAT TCC ATC	8 = 5				420 pro gin val ser thr pro thr leu val CCC CAA GTG TCA ACT CTA GTA	448 450 cys ala glu asp TGT GCA GAA GAC	\$67 7CC	S10 CAT
	A th	CAT	P 200	leu gly met TTG GGC ATG	thr leu glu lys ACT CTA GAG AAG	gin asn leu ile lys CAG AAT TTA ATC AAA	₽ 21 12	- 5	₹ ₹	ag 2
30	3 5	asn gln AAT CAA	glu met GAG ATG	500	66	¥ = .	570	448 pro eys CCC TCT	ŧδ	AC th
	ss A	glu asn GAA AAT	95	3 5	F C	. ₽ ¥	- 5 3	£ 55	573 573 705	를 E
	ACA th		\$ 2	를 E	th ACT	as AAT		ATG	476 477 cys cys thr TGC TGC ACA	glu thr phe thr
	4.010	265 978 1GT	glu asn GAA AAT	5 5 CTC	glu thr GAA ACC		Va1	Ę Ş	1ys	al &
35	lys leu AAG TTA	7 Te		asp CAT		5 5	* \$	₹ ₹	A th	als CCT
		7.4	291 ala glu val GCC GAA GTG	1ys	₹¥	381 val glu glu pro GTG GAA GAG CCT	411 tyr thr lys lys val TAC ACC AAG AAA GTA	۽ 5	471 asp arg val thr lys GAC ACA GTC ACC AAA	AAT A
	36.	261 ala lys GCC AAG	g SA	321 glu ala GAG GCA	351 1ys thr AAG ACA	58	i S	골중	45 P	å E
	2 2 5	261 818 GCC	29. 8.18 000	321 9 tu GAG	351 178 AAG	38 CFC	411 tyr TAC	# # CCT	471 88p	501 glu phe CAC TTT

5	550 SSS 559 540 SSS 559 540 SSS 559 540 SSS 559 SSS 55	561 560 Fer 570 term 567 570 term of the sale and ser gin ale ale ter giy tev ter 670 fer 570 see gin ale ale giy tev ter 670 fer 670 fer 670 fer 670 fer 670 fer 670 fer 770	ter ter Gatorogeacakoarangaamatoaaateaaagtaatteatteatteattittetttteettestestaarageareagtestaaaaataaaataaattaaa (2002)	
	0 5 X	9	ž	
	558 559 560 6ys cys lys TGC TGC AAG	± ₹	10.11	
	8 % SZ	ACA	AA11	
	AAG A	CATC	CATA	
10	216	5 ₹	¥	
	CTA 1	1 € T	1	
	å E	슬성	C1G1	
	1 5	3 €	CACC	
15	SCT S		Š	
13	4 5	£ 5	YYYC	
	550 CAT 750	95 E X3	50101	=
	es p	AG 25	E	202
	ATG	£ 55	Ĕ	₹.
20	val	£ 5	121	÷
	ala GCT	5 T	# 5	2
	¥, ₩	3 5	ATCT	÷
	55 CT 0	₹ ₹	ATC	Ĭ.
25	5₹	5. ₹	1130	AATC
25	95	영술	¥	XX
	AA.	골용	5	ATG
	8 t 8	55 g 85	Τζ¥	¥
	at s	د 8	¥.	AATA
30	1ys AAG	# E	AAAG	Y
	50	52 S	7 5	2110
	lys AAG	4 th	¥ 4	TCT
	5 A	48	ž ž	177
		¥ ¥	Ş	-
35	515	8 2	100	į
	3 5	\$ 3	5	**************************************
	53.4 9.10 646	25 to 50	ર્ટ	;

Nucleotide sequence coding for the pre pro human serum albumin, said nucleotide sequence is as follows:

5	(36)	(170)	(380)	(320)	(440)	(330)	(029)	. (917)	230 phe pro lys ala alu phe ala alu TTT CCC AAA GCT GAG TTT GCA GAA (300)
	- 1 8 8	lys AAA	8 = 5		5 5 5 5 c	1 t 1 t 2		200 1G 3	230 910 GAA
	₹ E					3 至		NG AG	2 5
	3 5	AAT	25	ala thr GCA ACT	asp asn GAC AAC	TAC	168 169 cys cys TGT TGC	3 8	å E
10	å E	5.5	¥ t		65 P	2, ₹	2 4 S	£ 5	2 %
10	3 5	-3 ≾	18 Z	i S	\$ \$	1, ¥	¥ \$	# 8	25
	F 3 5	충용	₹.3	£ \$ \$	leu gln his lys TTG CAA CAC AAA	3 2	168 ale ale phe thr alu cys GCT GCT TTT ACA GAA TGT	1,4 A	N A
	2 22	3 5	AAT	3 ¥	- £ 5	å E	5 to	<u>ون</u> ي	ę g
	# E	CA T	510	75 lys leu cys AAA TTA TCC	3 2	ž Ş	# L3	ser ala lys TCT GCC AAA	a E
15	a E	lys AAA	3 5	GAC	를 E	금왕	1 ys	36 100	P AG
	p r c trp val tlu phe lle ser leu leu phe leu bhe TGG GTA ACC TTT ATT TCC CTT CTT TTT CTC TTT	10 glu val als his ang phe lys aso leu qiy qlu glu san phe GAG GTT GCT CAT TCG TTT AAA GAT TTG GCA GAA AAT TTC	30 gin tyr leu gin gin cys pro phe diu aap his val lys leu val aan giu val thr diu bhe CAG TAT CTT CAG CAG TGT CCA TTT GAA GAT CAG ATA TTA GTG AAT GAA GAA TTA	50 62 as a ser leu his the leu dhe djy aso lys leu cys the val cc ca Ant TGT GAC ANT TGT G	100 101 qlu cys phe leu qln his lys GAA TGC TTG CAA CAC AAA	diu giu thr phe leu lys lys GAA GAG ACA TTT TTG AAA AAA	160 phe phe ala lys arg tyr lys ala ala phe thr TTC TTI GCT AAA AGG TAT AAA GCT GCT TTT ACA	€ 5	220 val ala arg leu ser gln arg GTA GCT CGC CTG AGC CAG AGA
	. , 5	5 5 5	5 to 45	S of F	울 물 중	130 AAT	8 F S	130 AAG	220 AGC
	를 물	£ £	5 5	3 5	AAT	130 asb ash GAC AAT	1ys AAA	£ 99	3 E
••	-18 Het lys ATG AAG	sls GCT	as CAT	k th	90 91 cys cys ata 1ys gin giu pro aly ara asn alu cys rcc rcr ccc AcA AAT CAA TGC	phe his TTT CAT	phe phe ala TTC TTT GCT	srg asp alu aly CGG GAT GAA GOG	5 8
20	4 F -	£ 5	8 8	5 5	90 91 aly giu met ala asp cya cya ala lya gin giu bro aly ara gat gwa atg cat cwt toc tot cax ana cna gna cat asa Asa	å E	a E	es CAT	동
			ďΕ	3 E	2 L3	÷ 5	a E	5.5	54 Y
•		AGT BE	e y	lys ser AAA TCA	6 8	Ł tr	3 €	glu leu GAA CTT	£ 59
		lys AAG	4 8 P	₹ }	둫	\$ \$ 5±	150 tyr ala pro glu leu leu TAT GCC CCG GAA CTC CTT	중중	210 ala phe lys ala trp ala GCT TTC AAA GCA TGG GCA
25		\$ 55 S	튭왕	gg Cyc	1, A	ATG T	5 5	\$ 2	* 55
		- 8	45	62 Cys	: 5	15.0	£ 20	≅ €	1ys
		- 5 7	₹ 5	60 62 glu asn cys asp GAA AAT TGT GAC	9 8 101	GAT SP	÷ 8	180 pro lys leu CCA AAG CTC	를 E
		-6 p r o -1 1 arg gly val phe arg arg asp ala his lys ser AGG GGT GTT CGT CGA GAT GCA CAC ANG AGT	30 gin tyr leu gin gin cys pro CAG TAT CTT CAG CAG TGT CCA	용물용	8 \$ 5	120 glu val asp val met cys thr CAG GTT GAT GTG ATG TGC ACT		\$ 50	210 818 GCT
		8.79 CG1	£ 95	818 GCT	\$ 5	£ %	arg his pro tyr phe AGA CAT CCT TAC TTT	177 ala cys leu leu GCC TGC CTG TTG	£ 2
30		ૄ૾ૠ૾ૼ E	ala phe ala OCC TTT CCT	ala asp glu ser GCT GAT GAG TCA	5 5	5 Y	his pro tyr CAT CCT TAC	≅ £	쿰죙
		7 1 2 2	# E	asp glu GAT GAG	glu met GAA ATG	val arg GTG AGA	5 5	E & B	58
		g 13	# 8	\$ 2	爰ჴ	575	₹ ₹	1 20	# F
		4 5 8	leu ile TTG ATT		aly cct	15 T	5 5	sls CCT	lys AA
35		-1 -6 pro -1 1 tyr ser arg gly val phe arg arg asp ala his lys ser rar rec Acc cor dic TTT cor coa CAT COA CAC ANG	21 ala leu val leu GCC TTG GTG TTG	53 cys val TGT GTT	81 arg glu thr tyr CGT GAA ACC TAT	120 Fro arg leu val ang pro glu val asp val met ops thr ala phe his asp aan glu glu thr nhe leu lya lya tyr leu CCC CCA TTG GTG ACA CCA CAG GTT GAT GTG ATG CCT TTT CAT GAC ANT GAA CAG ACA TTT TTG AAA AAA TAG TTA	141 glu fle ala arg GAA ATT GCC AGA	170 180 177 180 also are are also by a leu say giu leu are asp diu gip lys als ser ser als lys din are leu lys als als asp lys als also cys leu leu pro lys leu asp giu leu are asp diu gip lys als ser ser als lys din are leu lys 6cT 6cT 6cX 6cX 6cX 6cC 16c CGX 716c CGX 6cX 6cX 6cX 10cC 6cX 6cX 6cX 6cX 6cX 716cX 716cX 717 717 717 717 717	220 als ser leu gin lys phe gly glu arg als phe lys als trp als val als arg leu ser gin arg phe pro lys als glu phe als GCC AGT CTC CAA AAA TIT GCA GAA AGA GCT TTC AAA GCA TGG GCA GTA GCA CCC CGC AGC AGA TTT CCC AAAA GCT CAG TTT CCA
55		ty.	5 ST	≈ 85 ₽	A th		8 59 80	GAT A	15 C
		ser als TCG GCT	3 5	51 Iys thr AM ACA	₹ §	111 asn leu AAC CTC	# E	5 5	AGT
		700	2 د و	2 ₹ ≸	E 15	± 8 5 ₹	₹ ₹ §	£ # P	201 818 GCC

5	(890)	(980)	(1070)	(1160)	(1250)	(1340)	(1430)	(1520)	(1610)	(1700)
			320 81 8 0:1		CCT Lea	5 to 50		£70	Sho AAA	530 E 1
	£ 5 €	289 103	1 ty	350 leu aro leu ala CTG AGA CTT GCC	52	val GTT	440 1ys hls AAA CAT	15 YE	2 2	3 E
	- 8		lys asn tyr AAA AAC TAT	E Q	\$ ₹	leu leu CTG TTA	437 438 cys cys TGT TGT	6 A	ra ta	£ 59
10	5.5	ser his Tor CAC	310 ala asp leu pro aer leu ala ala asp phe val qlu aer lys asp val cys lys asn tyr ccr cot rrc ccr rch rrk ccr ccr cot mir cir cok kgr Acc Acc Acc Trc Akk Akc Tkr	340 glu tyr ala ang ang his pro asp tyr ser val val leu leu leu ana leu GAR TAT GCA AGA AGA GAT GAT GGT GTG GTG CTG CTA AGA CTT	369 370 ale ale asp pro his diu cys tyr ale lys wel phe asp diu phe lys pro GCT GCA GNT CCT CAT GCC AAA GTG TTG GNT GAA TTT AAA CCT	400 yin aan leu ile iya gin aan eya gitu leu phe cilu gin leu qiy giu tyr iya phe gin aan ala leu ieu wal OoG ANT TIA ATC AAN CAA AAT GCT GOLG CETT TIT GOLG CETT GGA GAG TAC AAT TIC CAG AAT GCT CTG TTA GTT	437 038 1GT	th ACG	476 477 480 especially assess and any any promouse peer also led glu wal asses old the tyr wall promise eys eys the glu see led wall asses any any promouse peer also led glu wal asses old the tyr wall promise. Too mote and the control of the cont	ACT TA
10	25 PB	<u>\$</u> ¥	316 cys	val val leu leu GTC GTG CTG CTG	370 tyr ala lys val phe asp olu TAT GCC AAA GTG TTC GAT GAA	- 8	1 ys	atu 1ys GAG AAA	¥ Ş	520 giu Iya giu ara din 11e Iya Iya ain. GaG AAG GAG AGA CAA ATC AAG AAA CAA
	8 2	큺죓	3 F	3 25	885 CAT	AAT AAT	Ser AGC	12 05	급성	1 ys
	= 5	3 E	\$\$ ₹	55	T S	f 95	val qly GTG GGC	₹ 5	\$ 5	AAG AAG
	240 255 255 256 250 250 250 250 250 250 250 250 250 250	278 279 280 glu cys cys glu lys pro leu leu glu lys GM TGC TGT GM AAA CCT CTG TTG GAA AAA	310 glu met pro ala asp leu pro ser leu ala ala asp phe val glu ser lya aso val cys GAG ATG CCT CAC TCA TTA CCT CCT CAT TTT GTT GAA AGT AAG AAG AAG AAG AAG A	ž 59	510 CTC	400 gin leu gly glu tyr lys phe gin CAG CTT GGA GAG TAC AAA TTC CAG	val GTG	leu his TTG CAT	575	11e
15	급성	85	AGT	# E	1ys	2 ₹	\$ ¥	c c	5 €	등중
	250 253 Leu leu plu cys CTG CTT GAA TGT	N. A.	골중	1 × 2	÷ 58	7.3	430 arg asn leu qly lys AGA AAC CTA GGA AAA	460 461 1eu cys TTA TGT	of CTC	£ 5
	23 CT CT C	8 2 38	5 te 7 TE	340 asp tyr GAT TAC	370 1ÅT	64 g 69	1 to 1	460 461 gln leu cys CAG TTA TGT	490 ala CCT	55 ± 55
	8 2	279 cys TGT	a E	Dr.	369 alu cys GAA TGC	충용	AK AK	46 95	P 25	A S
	gly asp GGA GAT	278 279 leu lys glu cys cys CTG AAG GAA TGC TGT	\$ 5	₹ }	급충	를	ž ž	AAC	a E	16. SA
20	# E	큺		ale arg arg his GCA AGA AGG CAT	pro his CCT CAT	F. 95	glu val ser GAG GTC TCA	3 5	5 5	thr leu ser ACA CTT TCT
	3 % 5	ser lys leu lys AGT AAA CTG AAG	et cc	£ 5	5 2	물용	2 E	val val GTG GTC	£ 55	E E
	245 246 0ys 0ys TGC TGC	1 te	3 ₹	2 5 5 5	ala ala asp. GCT GCA GAT	ag E	8 8	550	£ 3	후호
	- ₹ ₹	1. A	1 2	glu tyr GAA TAT	£ 5	3 5	2 E	F 55	819 AGG	¥ %5
25	245 his the glu oys CAC ACG GAA TGC	AGT 36	PT CCT	5 ₹	÷ 5	glu leu phe alu GAG CTT TTT GAG	- 3 F	asp tyr leu ser GAC TAT CTA TCC	as AC	514 asp 11e cys GAT ATA TGC
	- 5	5 55	3 5	ξž	ala SS	392 cys TGT	ACT TH	₹¥	150	충
	510	# £	g S	3 5	360 361 oys cys ala 1GC TGT GCC	AT AT	8 Y	\$ 8	15	£ 55
	2 ≥ ₹	270 gln asp ser CAA GAT TCG	8 ÷ 5	3 a E	360 100 100	390 CA	420 thr ACT	450 ala glu GCA GAA	480 TCC	510 thr phe his ACC TTC CAT
	leu thr CTT ACC	es t	5 55	met ATG	lys AAG	₹. ₹	val ser GTG TCA	- 3	중중	# 5
30	3 5	튭죙	glu met GAG ATG	9,50	58	11e	4a GTG	448 cys	rt S	
	val thr amp leu GTG ACA GAT CTT	265 cys glu asn TGT GAA AAT	5 5	330 val phe leu gly met phe leu GTC TTC TTG GGC ATG TTT TTG	360 thr thr leu glu lys gys ACC ACT CTA GAG AAG TGC	gin asn leu lle lys CAG AAT TTA ATC AAA	튭	met pro ATG CCC	476 477 cys cys TGC TGC	glu thr phe GAA ACA TTC
	Şξ	58	asn asp AAT CAT		rt t	AAT	val pro GTA CCC	ATG	476 cys	효호
	5 CT C		¥ ¥	5 5	A SP	45	g val		₹. ¥	급정
	3 5	11e	95	GA TA	9f 8	5 E	1 ys	8, ₹	AC th	al a
35		7 14	val glu GTG GAA	1ys	ty 1¥1	16 95	1ys	£ 55	CTC CTC	AAT AAT
	# 5 5	261 ala lys tyr lle GCC AAG TAT ATC	291 ala glu val GCC GAA GTG	et a	351 lys thr tyr AAG ACA TAT	58	420 tyr thr lys lys val pro gin val ser thr pro thr leu val 7AC ACC AAC AAA GTA CCC CAA GTG TCA ACT CCA ACT CTT GTA	58	471 asp arg val thr GAC AGA GTC ACC	Æ E
	5 5 5	261 813 GCC	£ 200	330 glu ala lya asp val phe leu gly met phe leu tyr GAG GCA AAG GAT GTC TTC TTG GGC ATG TTT TTG TAT	351 1ys	val GTG	14 4 5	# £ 500	£ 450	501 91u GAG

5				(1790)	567 570 glu thr cys phe ala glu glu dly lys lys leu val ala ale ser gln ala ala leu dly leu ter GAG ACC TGC TIT GCC GAG GAG GAT AAA CTT GTT GCT GCA AGT CAA GGT GCC TIX GGC TIX TAX CATCHCATITAAAG (1883)	
			99	ķ Ş	ž	
			558 559	5 5	₹ ₹	
			558	\$ 5	ACAT	
10				ž Š	25	
10				름용	₹ŧ	
				7 E	3 ₹	
				å E	충용	
				불	3 5	
15	•			g 5	- 8	
				E E	eg 55	
			8	2 2	88 £ 33	
				\$ 5	36 T	
				ATG T	- 5	
20				3 E	ect S	
				5 5 5	2 E	
				\$ ₹	3 E	
				₹ E	¥ ¥	
25				ธ์ รั	₹ ¥	
				38	69	
				\$ ₹	916	
			240	ξŞ	570 910 GAG	
				2 §	ور ور	
30				Ş.Ş	å E	
				ž č	567 cys TGC	
				¥ 5	ž Ş	- 3
			. :	ž ž	916	:
		•	٠.	val Iya hia Iya pro Iya ala thr Iya giu gin leu Iya ala val met asp asp phe ala ala phe val giu Iya cya cya. GTG AAA CAC AAC CAC AAC AAA AAA GAG CAA CTG AAA GCT GIT ATG GAT GAT TTC GCT GCT TTT GTA GAG AAG TGC TGC	asp asp lys g	•
35				5 5 2 5	S P	
				3 E	8 25	
			£ .	SAG CTC O	ala CCT	

ter ter cocciaccaticolatargamamamatgamagettattcatetettittettitteettgetgamagecaacaeeetettamamaatamattettam (2002)

TCATTTTGCCTCTTTTCTCTGTGCTTCAATTAAAAAATGGAAGAATCTAA..... 20AA (2078)

- 10. A nucleotide sequence according to any of claims 6 to 9, in essentially pure form.
- 11. A DNA transfer vector comprising a nucleotide sequence as defined in claim 5.
- 5 12. A DNA transfer vector according to claim 11, transferred to and replicated in a micro-organism.
 - 13. A DNA transfer vector according to claim 12, which is a plasmid.
- A DNA transfer vector according to claim 13,
 wherein the plasmid is pBR322 or YEp6.
 - 15. A process for preparing human serum albumin, which comprises culturing a micro-organism according to claim 5.
 - A DNA transfer vector according to any of
- 15 claims 12 to 14, or a process according to claim 15, wherein the micro-organism is a bacterium or yeast.17. A vector or process according to claim 16,
 - wherein the bacterium or yeast is <u>E. coli</u> or <u>Saccharomyces</u> <u>Cerevisiae</u>.

Restriction Endonuclesse Map of Human Serum Albumin cDNA Clones

